DOCUMENT RESUME

ED 397 116 TM 025 250

AUTHOR Stevens, Joseph J.; Clauser, Patricia
TITLE Longitudinal Examination of a Writing Portfolio and the ITBS.
PUB DATE Apr 96
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Tests; *Elementary School Students; Elementary Secondary Education; *Longitudinal Studies; *Portfolio Assessment; Portfolios (Background Materials); Predictive Validity; Scores; *Secondary School Students; State Programs; Test Results; Test Use; *Writing Evaluation
IDENTIFIERS *Iowa Tests of Basic Skills; United States (Southwest)

ABSTRACT

Assessment results were examined for 2,351 students in a large Southwestern school district over a 4-year period. Assessment in the first year consisted of the full battery of the Iowa Tests of Basic Skills (ITBS) administered to all third graders in the district. In the following year (grade four), the same students participated in a state-mandated Writing Portfolio Assessment (WPA) program. In the third year, the fifth-grade year, the ITBS was administered and in the fourth year (grade six) the writing portfolio was administered using new prompts. Patterns of relationships were studied both within and between assessment methods. Results showed moderate to high predictive validities for the ITBS and low predictive validities for the WPA. Application of a longitudinal structural equation model indicated that later achievement (either ITBS or WPA) was related to prior achievement as measured by the grade-three ITBS, but not as measured by grade four WPA. In fact, relatively little variance in student WPA scores was accounted for using information from other assessment measures or occasions. An appendix presents the grade four and grade six writing prompts. (Contains 1 figure, 4 tables, and 17 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 397 116

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY
JOSEPH J. STEVENS

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Longitudinal Examination of a Writing Portfolio and the ITBS

Joseph J. Stevens

Patricia Clauser

University of New Mexico

Paper presented at the annual meeting of the American Educational Research Association, New York, NY, April, 1996.

Address correspondence to: Joseph Stevens, 113 Simpson Hall, University of New Mexico, Albuquerque, NM 87131; (505)277-4203; e-mail: jstevens@unm.edu

TM025250

BEST COPY AVAILABLE

Abstract

We examined the assessment results for 2,351 students in a large Southwestern school district over a four-year period. Assessment in the first year consisted of the full battery of the ITBS administered to all third graders in the district. The following year (Grade 4), the same students participated in a state mandated Writing Portfolio Assessment (WPA) program. In the third year (Grade 5) the ITBS was administered and in the fourth year (Grade 6) the writing portfolio was administered using new prompts. Patterns of relationships were studied both within and between assessment methods. Results showed moderate to high predictive validities for the ITBS and low predictive validities for the WPA. Application of a longitudinal structural equation model indicated that later achievement (either ITBS or WPA) was related to prior achievement as measured by the Grade 3 ITBS, but not as measured by Grade 4 WPA. In fact, relatively little variance in student WPA scores was accounted for using information from other assessment measures or occasions.

Longitudinal Examination of a Writing Portfolio and the ITBS

Current interest in improving the authenticity or relevance of assessment has led to the development of numerous alternative assessment programs. These programs often seek to address and correct perceived failures of traditional item types and testing formats such as an emphasis on recognition and recall rather than production skills and a focus on processes that are not directly relevant to learning (Camp, 1993; Quellmalz, 1986). The desire for alternative methods has led to increases in the assessment of writing, often accomplished using portfolio methods. In 1992, for example, thirty-nine states assessed student writing (NCREL, 1993). Although writing portfolio methods are increasing in popularity, to date their measurement quality is largely unknown (Herman, Gearhart, & Baker, 1993).

Messick (1994) has pointed out that claims of authenticity and greater construct relevance of alternative assessments are best viewed as validity arguments that must be evaluated empirically. In previous studies (Stevens, 1995; Stevens & Clauser, 1995a; Stevens & Clauser, 1995b), we have begun to assess the internal, concurrent, and discriminant validity of language ability and achievement as represented by a traditional assessment instrument, the Iowa Tests of Basic Skills (ITBS), and an alternative assessment instrument, the New Mexico Writing Portfolio Assessment (WPA). We have found that a number of the properties that are intended as characteristics of the assessment instruments may not be supported by empirical evidence. For example, despite equivalent labeling and description of constructs, language abilities measured by the two alternative approaches to assessment (ITBS vs. WPA) are largely divergent (Stevens & Clauser, 1995a; Stevens & Clauser, 1995b).

The present study reports initial analyses examining the relationships among these instruments and assessment methods in a four-year longitudinal study. Our interests were in examining concurrent, discriminant, and predictive validity of both the Writing Portfolio Assessment and the traditional, limited-response ITBS. Our purpose was to address a number of issues including: 1) what are the predictive validities of the traditional and alternative assessments? 2) how do instrument subtests intercorrelate at each grade level? 3) how well can prior achievement using the "same-method" of assessment (i.e., ITBS to ITBS or WPA to WPA) predict later achievement, and 4) how well can prior achievement using a "different-method" of

assessment (i.e., ITBS to WPA or WPA to ITBS) predict later achievement?

Method

Instruments. The ITBS is one of the most widely used and accepted measures of student achievement (Lane, 1992; Linn, 1989). The ITBS is a standardized, norm-referenced instrument, composed of multiple-choice and other limited-response items. The ITBS Multilevel Battery, Form J (Hieronymus & Hoover, 1986) was administered. For the purposes of the present study only the six ITBS language subtests were relevant: 1) Vocabulary, 2) Reading, 3) Spelling, 4) Capitalization, 5) Punctuation, and 6) Language Use and Expression. In addition, the Language Total, a composite of the six language subtests, was used. Internal consistency (KR-20) reliability of the ITBS language subtests is reported as ranging from .86 to .93 and was .96 for the Language Total in Grade 3. Internal consistency reliabilities for the language subtests in Grade 6 ranged from .82 to .91 with a reliability of .96 for the Language Total (Hieronymus & Hoover, 1986).

The New Mexico Writing Portfolio Assessment (WPA) is a program first administered in the 1991-92 school year and designed to provide an environment in which writing is valued and integrated into classroom activities. The program is organized and administered by the State Department of Education. The assessment is mandated at the fourth and sixth grades and is optional at eighth grade. Scoring is accomplished by an out-of-state contractor.

The WPA is unusual because it involves prompts that are not secure. In the Fall, the state sends three prompts to the schools at each of the three participating grade levels. Students are asked to write to these prompts and to collect their writing in a portfolio. Emphasis is on regular practice and feedback on the components of the writing process. The teacher's assessment manual contains information on how to work with the practice prompts, what procedures to follow when working with the required prompt, the scoring rubrics, descriptions of each mode of discourse (narrative, expository, and descriptive) and writing samples at each rubric score point. In February the state notifies the school districts which one of the original three prompts will be used as the operational piece. Students copy their final response for this required prompt into an official four-page booklet which is collected by the teacher and sent to

the State Department of Education for collection and scoring by a contractor.

A range-finding committee then scores anchor papers using a six-point rubric included in the teachers' manual. The range finders evaluate actual assessment papers which subsequently are placed as samples in a guide used to train the out-of-state raters to evaluate papers in the same way as the in-state teachers. Approximately eighty papers per prompt are evaluated as benchmarks. The range-finding committee includes classroom teachers, members of the State Department of Education assessment and evaluation staff, and two representatives from the contractor. During actual scoring of the student papers, a holistic score is obtained from two readers with adjudication by a third reader when disagreements of two or more points on a six-point scale occur. Additionally, all responses are scored by a single reader on four analytic scales: 1) Development (i.e. organization, detail, and clarity of writing), 2) Word Usage (i.e., correct use of vocabulary and grammatical forms), 3) Sentence Formation (i.e., correct use of sentence structure), and 4) Language Mechanics (i.e., correct use of punctuation, capitalization, and spelling).

Reliability coefficients are not available for the analytic scores on the WPA which are read by one rater. For the Holistic scores, reliability was computed by determining the percentage of papers on which rater agreement was exact, differed by one point, or differed by two or more points on the six-point scale. For the 4th grade administration in 1993, there were 64% exact agreements, 34% 1-point disagreements, and 3% disagreements of 2 or more points. For the 6th grade administration in 1995, there were 58% exact agreements, 39% 1-point disagreements, and 4% disagreements of 2 or more points.

Sample and Procedure. Computerized records for the ITBS and the Writing Portfolio Assessment were collected and matched for all students in a large suburban school district in the Southwestern United States over a four year period from Spring, 1992 to Spring, 1995. Students took the ITBS Form J-multilevel battery (Hieronymus & Hoover, 1986) in the Spring of 1992 and the Spring of 1994. The Writing Portfolio Assessment was administered in the Spring of 1993 and the Spring of 1995. WPA prompts for both years are contained in the Appendix. In Grade 4 a descriptive essay prompt was used and in Grade 6 a narrative prompt was collected as the operational piece in the portfolio.

Students were matched on identification number, name, and self-reported ethnicity. Students were eliminated if they had not taken all assessments, or if there was a mismatch of name or self-reported ethnicity. Following matching and listwise deletion of incomplete cases, a sample of 2,351 students was obtained who had participated in all four assessment years.

Results

The first analysis conducted was an examination of the intercorrelations of scores for the ITBS and for the WPA over time. Table 1 shows the correlations for the ITBS subtests and for the Language Total. Listwise deletion of cases with missing data resulted in a total sample size of 2,338 for the data in Tables 1 and 2. Correlations listed above the diagonal in Table 1 are those from the third grade administration and correlations below the diagonal are from the fifth grade administration. Entries on the diagonal of Table 1 represent the intercorrelation of a subtest from third grade with the same subtest in the fifth grade. These correlations can therefore be interpreted as predictive validities. Intercorrelations of the ITBS scores were generally moderate to high at both grade levels. The average third grade correlation was .694 and the average fifth grade correlation was .714. These correlations support the interrelatedness of ITBS subtests at both grades as reported elsewhere (Klein, 1981; Martin & Dunbar, 1985; Stevens, 1995). ITBS predictive validities were generally high with an average correlation of .678 across the seven measures. Predictive validity was highest for the Language Total (.763) and was lowest for the Capitalization and Punctuation subtests (.559 and .588).

Table 2 shows the intercorrelations of the WPA scores. The upper diagonal of Table 2 shows correlations of scores in Grade 4 and the lower diagonal shows score correlations in Grade 6. As in Table 1, entries on the table diagonal represent predictive validities. Average correlations of the scores in Grade 4 was .507 and in Grade 6 was .570. Predictive validities for WPA scores averaged .241, markedly lower than those for the ITBS subtests. The highest WPA predictive validity was .301 for the holistic score and the lowest predictive validity was .188 for the Development score.

In addition to exploring predictive validity by subtest, we were interested in examining the pattern of relationships among the assessment devices over the four-year study interval. For

each year, the measures that represented those scores most likely to be used in a high stakes context were chosen for this stage of analysis. The Holistic score is emphasized in use of the WPA and is the only score that is double-read. This score was used to represent language achievement as measured by the WPA at fourth and sixth grades. The Language Total of the ITBS was chosen for further analysis as the logical choice if a single summary score of language achievement from the ITBS was used. Using these two measures, each administered at two points in time, later achievement was modeled using prior years of achievement as predictors. This resulted in the longitudinal model illustrated in Figure 1. The model is a just-identified structural equation model using only observed variables. As can be seen in Figure 1, initial language achievement is measured by the ITBS at Grade 3. This measure is then used to predict achievement at all three later grades. Language achievement as measured by the writing portfolio in Grade 4 is also used as a predictor of later achievement in grades 5 and 6, and lastly, the ITBS in Grade 5 is used as a predictor of achievement in Grade 6. Thus the model allows prediction of later language achievement based on earlier language achievement both within and across assessment methods.

Correlations, means, and standard deviations for the variables used in the structural equation model are listed in Table 3. Listwise deletion of cases with missing data resulted in a sample size of 2,351 for the data reported in Table 3 and Table 4. As can be seen in Table 3, variable intercorrelations were generally low to moderate in magnitude with the exception of the correlation between the two administrations of the ITBS (r = .763). While correlations are reported in Table 3 for ease of interpretation, the corresponding variances and covariances were used to examine the structural equation model in Figure 1. Model parameters were estimated using maximum likelihood methods as implemented in LISREL 8 (Jöreskog & Sörbom, 1993).

Results of the analysis are listed in Table 4 and in Figure 1. Parameter values in Figure 1 and below the diagonal of Table 4 are maximum likelihood estimates of the direct effects in the model. Where there are indirect effects in the model (e.g., from ITBS 3rd Grade to WPA 6th Grade through either WPA 4th Grade or ITBS 5th Grade) Table 4 also lists the total effects in the model. To obtain indirect effects, the parameter estimate for the direct effect can be subtracted from the estimate for the total effect. In addition, standard errors (in parentheses) and z-test values are listed below each parameter estimate in Table 4.

All parameters in the model were significant, although there was substantial variation in the magnitude of parameter estimates. Initial achievement at Grade 3 was predictive of achievement at all later grades. The strongest coefficient in the model was from ITBS Grade 3 to ITBS Grade 5 (.710). Although this path spanning a two-year interval was stronger than that for the one-year interval from ITBS Grade 3 to WPA Grade 4 (.370), this result was expected since it represented a "same-method" path. The coefficient from ITBS Grade 3 to WPA Grade 6, a three-year "different-method" path, was substantially smaller in magnitude (.134).

Surprisingly, a similar pattern of results was not observed for prediction of achievement using the Grade 4 WPA. The "same-method" path coefficient from WPA Grade 4 to WPA Grade 6, was small (.140) and essentially equal in magnitude to that for the one-year interval, "different-method" ITBS at Grade 5 (.144). The coefficient from ITBS Grade 5 to WPA Grade 6 ("different-method") was also noticeably larger in magnitude (.277) than those from WPA Grade 4. Thus, application of the structural equation model showed that, for both "same-method" *and* "different-method" paths, the ITBS Language Total scores were generally stronger predictors than the WPA Holistic score. Examination of the magnitude of the coefficients of determination (R^2_y) for the endogenous variables in the model also showed that the WPA assessments were generally unrelated to other variables in the model: .137 for WPA Grade 4, .600 for ITBS Grade 5, and .216 for WPA Grade 6. The small magnitude of R^2_y for the WPA measures suggests that these assessments are not predictable on the basis of the information included in the structural equation model (including the "same-method" relationship between WPA Grade 4 and WPA Grade 6).

Discussion

In previous studies we found that method of assessment accounted for a larger proportion of score variance than the constructs being measured (Stevens & Clauser, 1995a; Stevens & Clauser, 1995b). These results suggested that mode of assessment would be a primary determinant of the relationships among student performances over time. That is, "same-method" relationships should be stronger than "different-method" relationships, even if the time interval for the former was greater than the latter. These predictions were not entirely supported by the results of the present study. While predictive validity was high for

subtests of the ITBS over a two-year interval, predictive validity for both the Holistic and Analytic scores on the Writing Portfolio Assessment were low across the two-year interval. Thus, while there was evidence of temporal stability of the ITBS, there appeared to be substantially less concurrent or predictive validity of the WPA over time.

Application of the longitudinal structural equation model provided further insight into the relationships among the instruments over time. The Grade 3 administration of the ITBS provided significant prediction of later achievement at all succeeding grade levels (total effects were .370, .763, and .397, for Grades 4, 5, and 6, respectively). While all coefficients were significant, there were substantive differences in the magnitude of the "same-method" coefficients in comparison to the "different-method" coefficients.

Relationships of the WPA at Grade 4 to later achievement were also significant, but were small in magnitude (total effects of .144 and .180 for Grades 5 and 6, respectively). While a small magnitude for the "different-method" coefficient was expected, the small "same-method" coefficient was not. In fact, additional evidence provided by coefficients of determination suggested that the WPA assessments at each grade level were largely unrelated to other variables in the model and were characterized by large proportions of score variance that were unique to the particular assessment.

There are several potential explanations for these results. The results indicated that measurement as provided by the WPA Holistic score is somewhat idiosyncratic and not related to later assessment of writing ability using the same methods and procedures nor to assessment of related language abilities using different methods and procedures (i.e., ITBS). One potential explanation for the observed results is that the mode of writing (descriptive vs. narrative prompts) produced differences in the assessment of student writing proficiency across the two grades. While this explanation might result in some suppression of the strength of relationships, it does not reconcile the relative superiority of the ITBS over Grade 4 WPA in predicting Grade 6 WPA.

A second explanation for the observed results is that differences in reliability account for the lower coefficients associated with the WPA. Application of a correction for attenuation, however, demonstrated that such an explanation accounts for only a portion of the magnitude of the observed

coefficients (for example, with perfect reliability, the correlation between Grade 4 and Grade 6 WPA is estimated as increasing from .302 to .451).

Another potential explanation for study findings is that operational use of a single prompt in the writing portfolio produces an assessment that is highly task or prompt dependent and creates little generalizability to other tasks or prompts. Limitations on generalizability as a result of task, prompt, or item sampling have been reported by several others (Burger & Burger, 1993; Linn, 1993; Linn & Burton, 1994; Shavelson, Baxter, & Gao, 1993).

Messick (1994) also describes difficulties that may arise from a task- rather than a construct-centered approach to assessment, including limited coverage of the content domain and measurement of features of the task that are construct-irrelevant. While results of the present study are preliminary, it appears likely that at least some degree of the unrelatedness of one WPA scores to another or to ITBS scores are a function of task specific features of the WPA. This suggests at least one improvement in assessment procedures: the operational use of multiple prompts in the WPA to enhance content coverage and generalizability.

References

Burger, S., & Burger, D. (1994). Determining the validity of performance-based assessment. Educational Measurement: Issues and Practices, 13(1), 9-15.

Camp, R. (1993). The place of portfolios in our changing views of writing assessment. In R.E. Bennett & W.C. Ward (Eds.), Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment. Hillsdale, NJ: Lawrence Erlbaum Associates.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Herman, J.L., Gearhart, M., & Baker, E.L. (1993). Assessing writing portfolios: Issues in the validity and meaning of scores. Educational Assessment, 1(3), 201-224.

Hieronymus, A.N., & Hoover, H.D. (1986). Manual for school administrators, Levels 5-14, ITBS Forms G/H. Chicago: The Riverside Publishing Company.

Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8 user's reference guide. Chicago, IL: Scientific Software International, Inc.

Lane, S. (1992). Review of the Iowa Tests of Basic Skills, Form J. In J.J. Kramer & J.C. Conoley (Eds.), The Eleventh Mental Measurements Yearbook. Lincoln, NE: The University of Nebraska Press.

Linn, R.L. (1989). Review of the Iowa Tests of Basic Skills, Forms G and H. In J.C. Conoley & J.J. Kramer (Eds.), The Tenth Mental Measurements Yearbook. Lincoln, NE: The University of Nebraska Press.

Linn, R.L. (1993). Educational assessment: Expanded expectations and challenges. Educational Evaluation and Policy Analysis, 15(1), 1-16.

Linn, R.L., & Burton, E. (1994). Performance-based assessment: Implications of task specificity. Educational Measurement: Issues and Practice, 13(1), 5-15.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. Educational Researcher, 23, 13-23.

North Central Regional Policy Information Center. (1993). State student assessment program data base, 1992-1993. Oak Brook, IL: Council of Chief State School Officers.

Quellmalz, E. (1986). Writing skills assessment. In R.A. Berk (Ed.). Performance assessment: Methods and applications. Baltimore, MD: Johns Hopkins University Press.

Shavelson, R., Baxter, G., & Gao, X. (1993). Sampling variability of performance assessments. Journal of Educational Measurement, 30(3), 215-232.

Stevens, J.J. (1995). Confirmatory factor analysis of the Iowa Tests of Basic Skills. Structural Equation Modeling: A Multidisciplinary Journal, 2(3), 214-231.

Stevens, J.J., & Clauser, P. (1995a). Multitrait-Multimethod comparisons of a writing portfolio and the ITBS. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

Stevens, J.J., & Clauser, P. (1995b). Anglo and hispanic students' performance on the ITBS and a writing portfolio. Paper presented at the annual meeting of the Rocky Mountain Educational Research Association, Albuquerque, NM.

Appendix

1992-93 New Mexico Portfolio Writing Assessment
Grade 4 Required Descriptive Prompt

Think about a special event you have been to. This could be a fiesta, a holiday celebration, a party, or any other special event. Describe this event so that someone who was not there will know what it was like. You might want to include what you saw, heard, and smelled, and how you felt when you were there.

1994-95 New Mexico Portfolio Writing Assessment
Grade 6 Required Narrative Prompt

Many times we wonder how something happens or why it happens. People think up stories to explain why things happen in nature. Use your imagination and have fun writing a story for your friends about one of the topics mentioned below. Choose one of the following "happenings" or pick one of your own and write a story to explain how it came to be.

How people came to have wrinkles

How cats came to have nine lives

How leopards came to have spots

How tears came to be salty

How giraffes came to have long necks

How the sea became salty

TABLE 1

ITBS Subscale Correlations and Predictive Validities
for Third and Fifth Grades (N = 2338)

	ITBS Score						
	1	2	3	4	5	6	7
1. Vocabulary	.721	.783	.647	.580	.561	.700	.730
2. Reading	.778	.725	.636	.602	.584	.715	.744
3. Spelling	.644	.655	.752	.638	.604	.657	.849
4. Capitalization	.568	.605	.626	.559	.694	.646	.862
5. Punctuation	.619	.645	.665	.725	.588	.632	.847
6. Usage/Expression	.724	.752	.656	.659	.693	.636	.864
7. Language Total	.738	.767	.854	.862	.886	.868	.763

Note. Correlations above the diagonal are for the third grade; correlations below the diagonal are for the fifth grade; underlined entries on the diagonal are predictive validities.

TABLE 2

WPA Score Correlations and Predictive Validities
for Fourth and Sixth Grades (N = 2338)

	WPA Score				
	1	2	3	4	5
1. Development	.188	.576	.530	.420	.558
2. Word Usage	.625	.212	.586	.507	.476
3. Sentence Formation	.555	.608	.250	.593	.462
4. Mechanics	.511	.546	.642	.256	.360
5. Holistic	.609	.553	.541	.508	.301

Note. Correlations above the diagonal are for the fourth grade; correlations below the diagonal are for the sixth grade; underlined entries on the diagonal are predictive validities.

TABLE 3
Correlations, Means, and Standard Deviations for Variables in the Longitudinal Structural Equation Model (N = 2351)

	Variable 1	2	3	4	Mean	SD
1. ITBS 3rd Grade	1.000				105.231	14.751
2. WPA 4th Grade	.370	1.000			2.633	0.717
3. ITBS 5th Grade	.763	.407	1.000		130.535	16.610
4. WPA 6th Grade	.397	.302	.436	1.000	3.163	1.036

TABLE 4

Direct Effects, Total Effects, Standard Errors, and *z*-Values for the Longitudinal Structural Equation Model

	Variable ITBS 3rd	WPA 4th	ITBS 5th	WPA 6th
ITBS 3rd			$.763(\gamma_{21})$	$.397(\gamma_{31})$
			(.013)	(.019)
			57.209	20.964
WPA 4th	$.370(\gamma_{11})$			$.180(\beta_{31})$
	(.019)			(.020)
	19.302			8.966
ITBS 5th	$.710(\gamma_{21})$	$.144(\beta_{21})$		
	(.014)	(.014)		
	50.527	10.288		
WPA 6th	$.134(\gamma_{31})$	$.140(\beta_{31})$	$.277(\beta_{32})$	
	(.028)	(.020)	(.029)	
	4.720	6.951	9.582	
Uniqueness		$.863(\Psi_{11})$	$.400(\Psi_{22})$	$.784(\Psi_{33})$
		(.025)	(.012)	(.023)
		34.520	33.333	34.087

Note. Direct effects are listed below the diagonal. When indirect effects are present, total effects are listed above the diagonal. Standard errors are listed in parentheses with *z*-test values below.

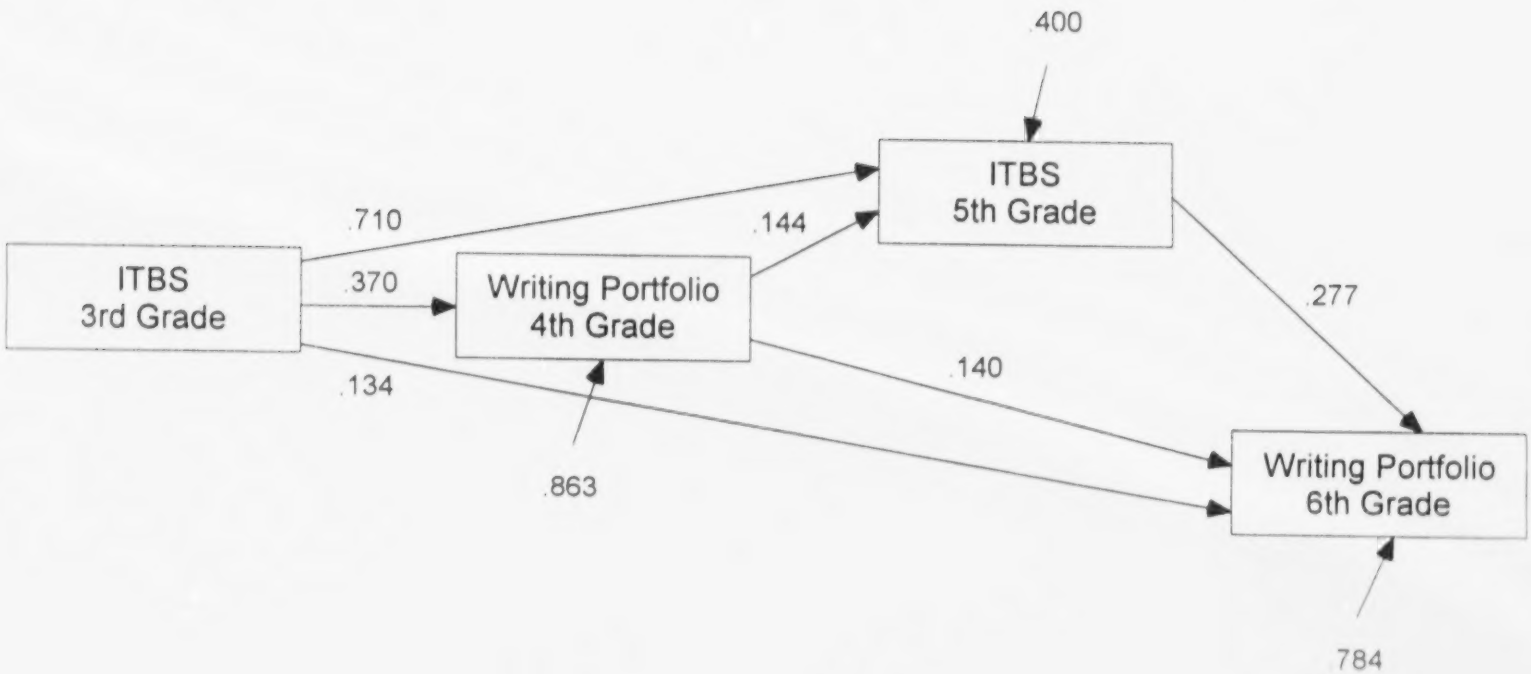

FIGURE 1 Longitudinal model of ITBS and Writing Portfolio scores over four years.